Investigating temporal changes of impression formed by interviewers from the applicant's behavior

Shota HIOKI, Michiko INOUE, and
Masashi NISHIYAMA[0000−0002−5964−3209]

Graduate School of Engineering, Tottori University
101 Minami 4-chome, Koyama-cho, Tottori, 680-8550, Japan
nishiyama@tottori-u.ac.jp

Abstract. In existing analytical studies on impressions in interviews, researchers have focused on subjective assessment after viewing stimulus video sequences and have not addressed the temporal changes of impression formed by interviewers. In this paper, we focus on the impressions that interviewers form from the applicant's behavior and conduct a sequential evaluation to capture the temporal changes of impression that evolve from moment to moment. For this evaluation, we built a system in a VR environment that can record in real time whether the impression is good, neutral, or poor, while keeping participant burden low. Using this system, we recorded the temporal changes of impression for 40 stimulus video sequences from 20 participants. We then computed inter-rater reliability for the temporal changes of impression across participants using the subjective scores recorded at each time point across all stimulus video sequences. The experimental results indicated that inter-rater reliability for the temporal changes of impression during interviews was moderate in the sequential evaluation.

Keywords: Interview · Behavior · Impression · Temporal changes · Sequential evaluation.

1 Introduction

In interviews conducted for recruitment and entrance examinations, the impression that an interviewer forms of an applicant is strongly influenced by the applicant's behavior [2, 9], which is part of the applicant's nonverbal information. Examples of the applicant's behavior include making hand and body gestures, establishing eye contact, and nodding. Figure 1 shows examples of such behavior. Numerous analytical studies have examined the applicant's behavior in interviews and the impressions that interviewers form of that behavior [7, 5, 10, 11]. Previous studies have typically investigated impressions related to social skills, such as sincerity, proactivity, cooperativeness, and initiative. They have also reported that interviewers' impressions vary with the applicant's behavior.

Fig. 1. Examples of behaviors performed by applicants in interviews.

In existing analytical studies [7, 5, 10, 11], statistical tests have been employed to investigate impressions formed from the applicant's behavior during interviews. As an initial step, it is common to examine inter-rater reliability, which measures the extent to which the impression formed from the applicant's behavior is shared across multiple participants who simulate interviewers. Previous studies have reported statistically significant inter-rater reliability across participants. In addition, these studies have used stimulus video sequences that simulate interviews to control experimental conditions. These studies evaluate impressions at the moment when each stimulus video sequence ends, and thus can be regarded as investigating the impression formed over the entire time span of the stimulus video sequence. In what follows, we refer to this conventional subjective assessment procedure as post-viewing evaluation of impression.

Because the timing at which impressions are formed is not exactly the same across participants, investigating how impressions change from moment to moment is important. However, existing analytical studies [7, 5, 10, 11] have not addressed how the impressions formed by participants who simulate interviewers change from moment to moment. To investigate such moment-to-moment changes, we focus on the time series of impressions recorded at each time point of a stimulus video sequence for each participant, which we refer to as the temporal changes of impression. In this paper, as a first step toward investigating the temporal changes of impression, we address the following two issues.

- **Issue** I_1: It is unclear to what extent the temporal changes of impression are shared across multiple participants.
- **Issue** I_2: There is no system that can record the temporal changes of impression while keeping participant burden low.

To record the temporal changes of impression, we consider a system that allows participants who simulate interviewers to perform subjective assessment in real time. While viewing a stimulus video sequence, participants are required to immediately input into the system the impression they hold at each time point. In what follows, we refer to this procedure for reporting the impression of a stimulus video sequence at each time point during viewing as sequential evaluation. Although this is not the objective of the present study, systems have been developed in the field of affective computing to record human emotions

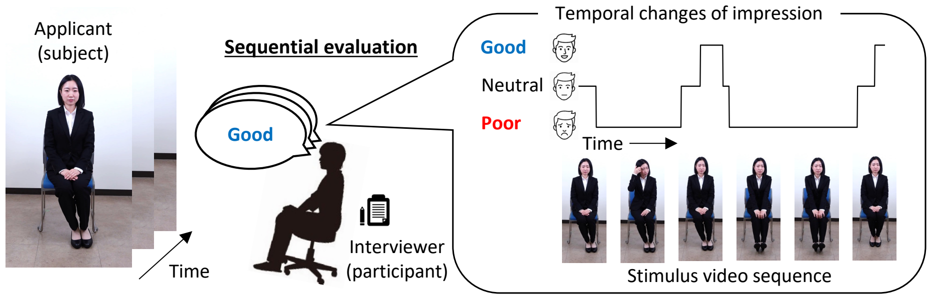

Fig. 2. Overview of the investigation in this paper: participants who simulated interviewers report good, neutral, or poor while viewing a stimulus video sequence in which a subject simulating an applicant exhibits interview behaviors, and the system records the temporal changes of impression.

from moment to moment in order to investigate their dynamic changes. Representative examples include FEELTRACE [4] and AffectTracker [6]. However, these systems are designed for recording emotion and are not suitable for investigating the impression that interviewers form from the applicant's behavior during interviews.

In this paper, we design a system to record the temporal changes of impression and investigate the extent to which the impression formed from the applicant's behavior is consistent over time across multiple interviewers. Figure 2 provides an overview of the investigation addressed in this paper. Although interviewers form various impressions, such as sincerity and proactivity, we focus on the overall impression related to social skills. We represent the impression using three levels: good, neutral, and poor. We assume a scene in which a subject simulating an applicant listens to the interviewer, and present a stimulus video sequence of several tens of seconds recorded from the subject. During viewing of the stimulus video sequence, participants who simulate interviewers report the impression in real time. The contributions of this paper are summarized as follows.

- **Contribution 1**: In the sequential evaluation of the temporal changes of impression, we performed a statistical test of inter-rater reliability and confirmed that inter-rater reliability was moderate across participants.
- **Contribution 2**: We developed a system that records the temporal changes of impression by allowing participants who simulate interviewers to report in real time whether the impression is good, neutral, or poor while viewing stimulus video sequences displayed on a VR (Virtual Reality) headset.

The experimental results showed moderate inter-rater reliability across participants for the temporal changes of impression in the sequential evaluation, although inter-rater reliability was lower than that in the post-viewing evaluation. Using the developed recording system for the temporal changes of impression, we confirmed that sequential evaluation can be performed by reporting impressions in real time while keeping participant burden low.

2 Related work

2.1 Effects of applicant behavior on interviewers' impressions

As described in Section 1, numerous analytical studies [7, 5, 10, 11] have investigated the impressions that interviewers form from the applicant's behavior in interviews. These studies have examined how the applicant's behavior affects impressions related to social skills. In these studies, post-viewing evaluation has been used to assess the impression formed over the entire time span of a stimulus video sequence. McGovern et al. [10] showed that applicants' nonverbal behavior, such as eye-contact and gesturing, tends to elicit impressions of confidence and proactivity. Gifford et al. [7] and Peeters et al. [11] showed that applicants' gesturing behavior and nodding behavior influence impressions related to social skills and cooperativeness. Furthermore, DeGroot et al. [5] revealed that interviewers rely strongly on applicants' gesturing behavior and postural behavior when forming impressions of the applicant's sincerity.

These existing analytical studies first confirm that inter-rater reliability in the post-viewing evaluation has often been reported to be high. To quantify inter-rater reliability, the intraclass correlation coefficient (ICC) is used. According to the guideline by Koo and Li [8], ICC values are interpreted as follows: less than 0.50 indicates poor reliability, 0.50 to less than 0.75 indicates moderate reliability, 0.75 to less than 0.90 indicates good reliability, and 0.90 or higher indicates excellent reliability. Gifford et al. [7] reported excellent inter-rater reliability on impressions related to social skills. Peeters et al. [11] reported good inter-rater reliability on the overall evaluation impression in interviews. DeGroot et al. [5] reported good inter-rater reliability on impressions of visual cues and personality traits.

Because these existing analytical studies use ICC based on post-viewing evaluation, they have not considered capturing the process itself in which impressions change from moment to moment, as described in Section 1. In this paper, by conducting a sequential evaluation for the temporal changes of impression, we aim to investigate inter-rater reliability using ICC.

2.2 Systems for recording subjective states in real time

To address Issues I_1 and I_2 described in Section 1, a system is required to perform sequential evaluation that handles impressions at each time point of a stimulus video sequence. In the field of affective computing, systems have been developed to record participants' subjective states in real time. For example, Cowie et al. [4] developed FEELTRACE. In this system, participants observe a stimulus video sequence on a stationary display and use a mouse or a handheld controller to record emotions that change from moment to moment. Participants input their emotions on a two-dimensional plane of valence and arousal. Fourcade et al. [6] extended this approach and developed AffectTracker, which can record emotions in real time in a VR environment. These systems enable real-time recording of the subjective state at each time point while viewing a stimulus video sequence.

However, in these existing systems [4, 6], participants must continuously and finely adjust subjective values in real time using a mouse or a handheld controller. Because this adjustment task imposes a substantial cognitive burden on participants, we consider it difficult to apply these systems directly to our problem setting. In this paper, we aim to develop a system that records in real time the impression that interviewers form from the applicant's behavior while reducing participants' cognitive burden as much as possible.

2.3 Duration of stimulus video sequences

The stimulus video sequences used in existing analytical studies [7, 10, 5, 11] can be broadly categorized into two types. In Type 1 [7, 10], the entire interview is recorded from start to finish, whereas in Type 2 [5, 11], only a segment within the interview is partially recorded. The duration of Type 1 stimulus video sequences is several tens of minutes, while that of Type 2 stimulus video sequences ranges from several seconds to several minutes. Because Type 1 requires substantial time and effort for subjective assessment, it is desirable to reduce participant burden. Type 2 uses short stimulus video sequences known as Thin Slices [1], which alleviates participant burden. However, regardless of which type is used for subjective assessment, impressions are evaluated at the moment when the stimulus video sequence ends, and thus sequential evaluation is not performed. In this paper, we aim to perform sequential evaluation while keeping participant burden low by using Type 2 stimulus video sequences.

3 Overview of the sequential evaluation of impression

3.1 Assumed scene

We discuss the scene addressed in this paper. In many interviews, a conversation between an interviewer and an applicant is conducted while both are seated. During the conversation, the interviewer speaks to the applicant and the applicant speaks to the interviewer in an alternating manner. In this paper, we focus on the scene in which the interviewer speaks to the applicant and the applicant listens to the interviewer. The interviewer is assumed to be seated at a position where the applicant's full body is visible. In this scene, the interviewer forms impressions from the applicant's behavior. In addition, the impression formed by the interviewer changes from moment to moment in accordance with the applicant's behavior.

3.2 Design policy for Issue I_1

As a design policy for addressing Issue I_1 described in Section 1, we investigate inter-rater reliability for the temporal changes of impression in the sequential

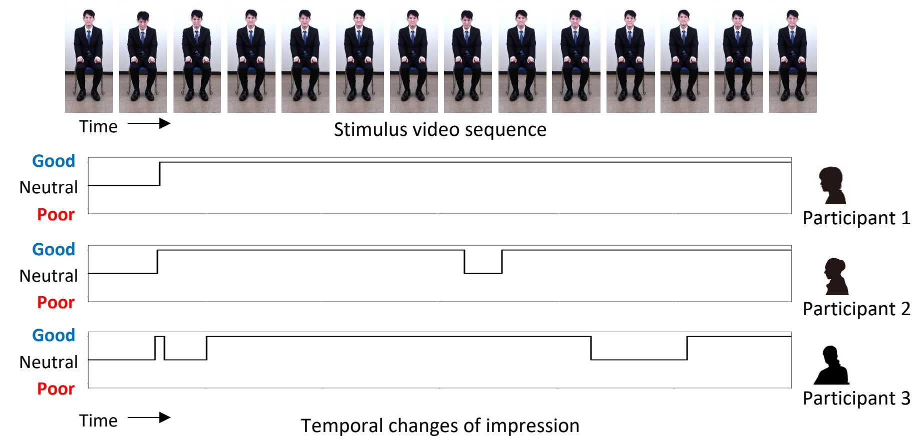

Fig. 3. Example graphs of the temporal changes of impression formed by each participant who simulated an interviewer while viewing a stimulus video sequence that includes the applicant's behavior, such as nodding.

evaluation. Specifically, we recruited multiple participants who simulated interviewers and tested the following hypothesis.

- **Hypothesis** H: The temporal changes of impression obtained by the sequential evaluation show statistically significant inter-rater reliability across participants who simulated interviewers.

Figure 3 shows an example of a stimulus video sequence together with a graph of the temporal changes of impression recorded by our system. The figure illustrates that, as time progresses in the stimulus video sequence and the applicant's behavior changes, the impression formed by a participant also changes over time in accordance with the behavior. In the figure, some participants (e.g., Participant 1 and Participant 2) show similar temporal changes of impression, whereas others (e.g., Participant 3) show temporal changes of impression that differ from those of other participants. To investigate inter-rater reliability for the temporal changes of impression, we compute intraclass correlation coefficients (ICC) [8].

3.3 Design policy for Issue I_2

As a design policy for addressing Issue I_2 described in Section 1, we develop a system that allows participants to report the impression at each time point of a stimulus video sequence. In our system, we provide participants with an impression question in advance. While participants view the stimulus video sequence using a head-mounted display (HMD), they use a handheld controller to report the impression quickly and intuitively. Figure 4 shows an overview of the system that conducts sequential evaluation to record the temporal changes of impression.

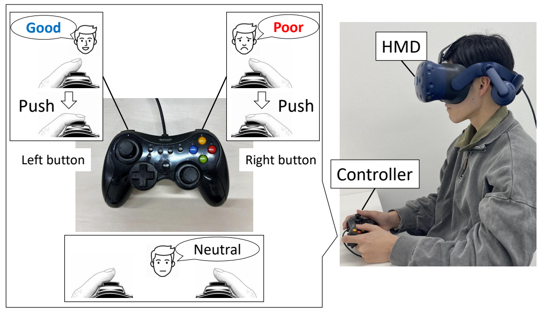

Fig. 4. Overview of the system that conducts sequential evaluation to record the temporal changes of impression.

Here, we discuss a specific approach for reporting impressions. A commonly used approach for reporting impressions is a multi-level scale such as a five-point Likert scale. However, reporting impressions on such a multi-level scale while viewing a stimulus video sequence would impose a substantial cognitive burden on participants. Therefore, in this paper, we aim to reduce cognitive burden by using a three-level scale with a handheld controller.

4 Experimental conditions of the sequential evaluation

4.1 Questions provided to participants

Twenty participants who simulated interviewers (Japanese; 15 male and 5 female; mean age: 22.6 $\pm$ 2.7 years) participated in the sequential evaluation.

We asked participants to report their impressions while assuming that they were responsible for a recruitment interview. When participants observed the subject simulating an applicant in the stimulus video sequence, we provided them with the following question.

Q: At each time point, please select one of the following three options to report the impression you form of the applicant.

- Good
- Neutral
- Poor

Participants were instructed to answer Question Q at all time points of each stimulus video sequence. Participants entered their responses using the handheld controller described in Section 3.3. Specifically, they pressed the LB button when the impression was good, pressed the RB button when the impression was poor, and released both buttons when the impression was neutral. To help participants understand which option they were inputting, visual feedback was displayed on the virtual screen. Each option was represented by the color of the frame surrounding the stimulus video sequence.

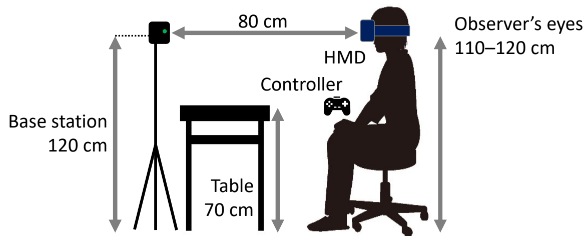

Fig. 5. Setup for our investigation.

4.2 Stimulus video sequences

We recorded stimulus video sequences assuming the scene described in Section 3.1, in which an applicant listens to the interviewer. The subjects in the stimulus video sequences were 20 university students simulating applicants (10 male and 10 female). For each subject, we recorded a video sequence S_g in which the subject exhibited behavior intended to give a good impression and a video sequence S_b in which the subject exhibited behavior intended to give a poor impression. The duration of each video sequence was 30 seconds. We recorded 40 stimulus video sequences (20 subjects $\times$ 2 conditions: good and poor).

For each participant, we used 36 stimulus video sequences randomly selected from the 40 recorded sequences. Specifically, we randomly selected 18 stimulus video sequences from S_g, consisting of 9 male and 9 female subjects. We refer to the selected stimulus video sequences intended to give a good impression as $\tilde{S}_g$. Using the same procedure, we selected 18 stimulus video sequences from the 20 sequences in S_b. We refer to the selected stimulus video sequences intended to give a poor impression as $\tilde{S}_b$. To analyze inter-rater reliability, we designed the experiment so that each stimulus video sequence was evaluated by multiple participants, ranging from 17 to 19 participants per sequence.

4.3 Setup

Figure 5 shows the setup for the sequential evaluation. To allow participants to view the stimulus video sequence, we used a head-mounted display (HMD; VIVE Pro Eye). The stimulus video sequence was presented on a two-dimensional virtual display in the VR environment. The refresh rate of the VR environment was set to 90 Hz, and the frame rate of the stimulus video sequence was set to 30 fps. To answer Question Q described in Section 4.1, participants held a handheld controller (ELECOM JC-U3613M). The sampling rate for recording impressions via the controller was set to 90 Hz.

4.4 Procedure of the sequential evaluation

As a preliminary step, we explained to participants the question and response approach described in Section 4.1. We then conducted the following steps.

P_1: We randomly selected one stimulus video sequence without replacement from the 36 stimulus video sequences $\tilde{S}_g$ and $\tilde{S}_b$ described in Section 4.2.
P_2: We placed a cross mark at a random position on the virtual display and asked the participant to fixate on it for approximately 2 seconds.
P_3: We presented the stimulus video sequence selected in Step P_1 on the virtual display and played it from start to end.
P_4: While viewing the stimulus video sequence, the participant answered Question Q in Section 4.1 using the handheld controller.
P_5: We provided the visual feedback described in Section 4.1 based on the participant's response.
P_6: We continued Steps P_4 and P_5 until the stimulus video sequence presented in Step P_3 reached its end.
P_7: We repeated Steps P_1 to P_6 until the participant completed reporting impressions for all stimulus video sequences in $\tilde{S}_g$ and $\tilde{S}_b$.

To avoid center bias [3], we introduced Step P_2. In the system implementation, we ran the process that displays the stimulus video sequence and the process that acquires participants' responses in parallel using separate threads, and recorded the responses by synchronizing them with the time points of the stimulus video sequence.

4.5 Computation of inter-rater reliability

To analyze inter-rater reliability for the temporal changes of impression, we converted the response options recorded in Step P_4 in Section 4.4 into numerical subjective scores. Specifically, we converted the good option to 1, the neutral option to 0, and the poor option to -1. For the analysis of inter-rater reliability, we computed ICC(2,1) and interpreted its magnitude according to the guideline by Koo and Li [8]. We computed ICC(2,1) using a two-way random-effects, absolute-agreement, single-measurement model, treating both raters (participants) and rating targets (subjective scores at each time point of each stimulus video sequence) as random effects. For implementation, we used the `rptR` package in the statistical computing environment R.

4.6 Post-viewing evaluation for comparison

To compare the results of the sequential evaluation with those of the post-viewing evaluation, we conducted the following procedure after Step P_6 in Section 4.4 was completed. Specifically, as a post-viewing evaluation, participants answered the following questions using a five-point Likert scale.

$\tilde{Q}_1$: Did you form a good impression of the applicant?
$\tilde{Q}_2$: Did you form a poor impression of the applicant?

For each question, participants selected one of five scores (5: strongly agree, 4: agree, 3: neutral, 2: disagree, 1: strongly disagree). When computing ICC(2,

Table 1. Inter-rater reliability (ICC) for the temporal changes of impression formed from the applicant's behavior by participants who simulated interviewers, in the sequential evaluation and the post-viewing evaluation.

	ICC	95% CI		p-value	Interpretation of ICC
		Lower	Upper		
Sequential evaluation	0.568	0.553	0.578	$< .001$	Moderate
Post-viewing evaluation	0.885	0.817	0.923	$< .001$	Good

1), we converted the five-point scores in the post-viewing evaluation into three-level scores comparable to those in the sequential evaluation. Specifically, we subtracted the score of $\tilde{Q}_2$ from the score of $\tilde{Q}_1$. For the resulting value, we assigned 1 when it was between 2 and 4, 0 when it was between -1 and 1, and -1 when it was between -4 and -2. Except for the conditions described above, the experimental conditions were the same as those for the sequential evaluation.

5 Experimental results

5.1 Results of inter-rater reliability

Table 1 shows inter-rater reliability for the temporal changes of impression formed by interviewers from the applicant's behavior in the sequential evaluation and the post-viewing evaluation.

First, we describe the results of the sequential evaluation newly addressed in this paper. We applied the experimental conditions for the sequential evaluation described in Sections 4.1 to 4.5. Inter-rater reliability in the sequential evaluation was ICC(2,1) = 0.568, 95% CI [0.553, 0.578], $p < .001$. This indicates statistically significant moderate reliability.

Next, we describe the results obtained using the post-viewing evaluation adopted in existing analytical studies. We applied the experimental conditions for the post-viewing evaluation described in Section 4.6. Inter-rater reliability in the post-viewing evaluation was ICC(2,1) = 0.885, 95% CI [0.817, 0.923], $p < .001$. This indicates statistically significant good reliability. This finding is in line with prior analytical studies [7, 5, 11] reporting good inter-rater reliability in the post-viewing evaluation, although their stimuli, target impressions, and experimental procedures differ from ours.

Overall, although inter-rater reliability in the sequential evaluation was lower than that in the post-viewing evaluation, the temporal changes of impression formed from the applicant's behavior showed moderate inter-rater reliability across participants who simulated interviewers.

5.2 Clustering of the temporal changes of impression

Purpose of clustering In Section 5.1, inter-rater reliability is summarized as a single value, and thus it does not explicitly reveal how impressions change over time. To examine such temporal patterns, we consider applying clustering to the temporal changes of impression of individual participants. By assigning each participant to a cluster, we aim to capture dominant patterns of the temporal changes of impression within each cluster. We focus on the number of participants assigned to each cluster to identify dominant patterns. For each cluster, we define a representative temporal pattern as the one that is closest to the cluster mean, while prioritizing clusters with larger numbers of participants.

Approach When performing clustering, a distance metric is required to measure the similarity of the temporal changes of impression across participants. We computed pairwise Manhattan distances between participants' temporal changes of impression and performed hierarchical clustering based on these distances. In this experiment, we empirically set the number of clusters to four. From the 40 stimulus video sequences, we selected the stimulus video sequence with the highest total subjective score across time points in the sequential evaluation and the one with the lowest total subjective score. For each selected stimulus video sequence, the number of participants was 18.

Clustering results Figure 6 shows the clustering results for the stimulus video sequence with the highest total subjective score in the sequential evaluation. Among the 18 participants, 9 were assigned to Cluster 1, 5 to Cluster 2, 2 to Cluster 3, and 2 to Cluster 4. Here, we discuss the patterns observed in Cluster 1 and Cluster 2, which included larger numbers of participants. In the representative temporal pattern of Cluster 1, participants formed a good impression from the beginning and maintained a good impression until the end. Specifically, around 3 seconds into the stimulus video sequence, participants formed a good impression in accordance with the nodding behavior, and this impression persisted until the end. This suggests that once a good impression was formed, it continued regardless of whether the behavior was present. In the representative temporal pattern of Cluster 2, participants formed a good impression intermittently, responding to each nodding behavior of the applicant. Specifically, similar to Cluster 1, participants formed a good impression around 3 seconds, but it soon returned to neutral. After that, participants formed a good impression each time the applicant nodded.

Figure 7 shows the clustering results for the stimulus video sequence with the lowest total subjective score in the sequential evaluation. Among the 18 participants, 13 were assigned to Cluster 1, 3 to Cluster 2, 1 to Cluster 3, and 1 to Cluster 4. Here, we discuss the pattern observed in Cluster 1, which included the largest number of participants. In the representative temporal pattern of Cluster 1, participants formed a poor impression from the beginning and maintained a poor impression until the end. Specifically, around 4 seconds into the stimulus video sequence, participants formed a poor impression in accordance with the applicant's bowing-head behavior, and this impression persisted until the end. This suggests that once a poor impression was formed, participants continued to hold the poor impression thereafter.

Table 2. Representative comments from participants about the system for recording the temporal changes of impression.

Category	Comment
Usability	
	A1: I quickly became familiar with which controller button corresponded to a good impression and which corresponded to a poor impression.
	A2: It was easy because I simply reported my impressions intuitively using the buttons, rather than thinking deeply.
	A3: Because it was my first time holding a game controller, I was slightly confused about the operation.
Burden	
	A4: Because the video sequence did not stop partway through, I could stay focused and report my impressions as natural reactions.
	A5: I did not feel any particular burden compared with the post-viewing evaluation, and I felt that I could continue even during a long experiment.
	A6: The HMD was heavy, and my head became tired when using it for a long time.

These results indicate that several patterns can be observed in the temporal changes of impression. Specifically, we observed a pattern of maintaining an impression until the end once it was formed from the applicant's behavior, as well as a pattern of forming impressions intermittently by responding to the behavior each time it occurred.

5.3 Qualitative evaluation of the developed system

We qualitatively evaluated the extent to which our system for recording the temporal changes of impression addressed Issue I_2 described in Section 1. Specifically, we conducted a questionnaire survey with participants after the sequential evaluation and the post-viewing evaluation were completed. We asked participants to provide free-text comments on the system's usability and perceived burden.

Table 2 summarizes representative comments from participants about the developed system. As in A1, A2, A4, and A5 in the table, we obtained many positive comments regarding usability and perceived burden. We also obtained a small number of negative comments, such as A3 and A6. We consider that our system has a high potential to record the temporal changes of impression without imposing excessive burden on many participants.

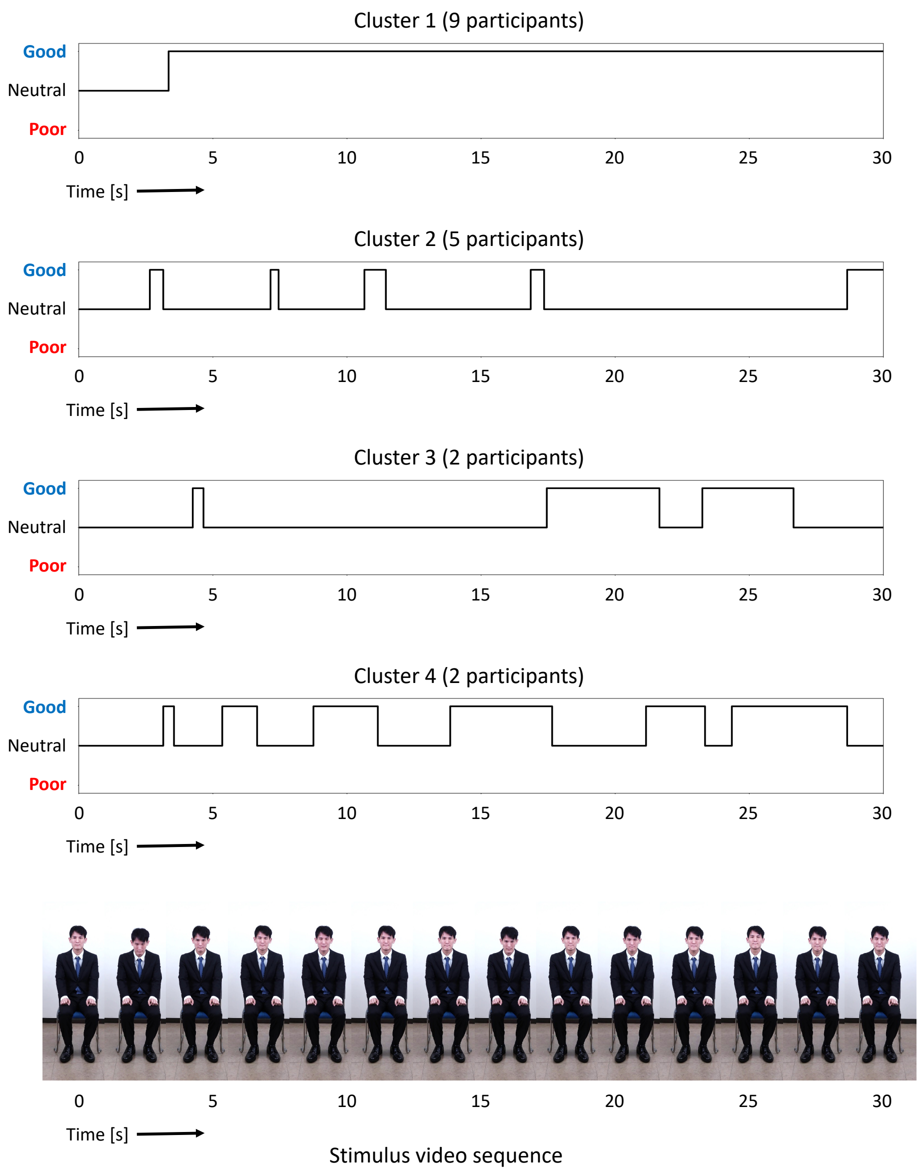

Fig. 6. Clustering results of the temporal changes of impression for the stimulus video sequence with the *highest* total subjective score in the sequential evaluation.

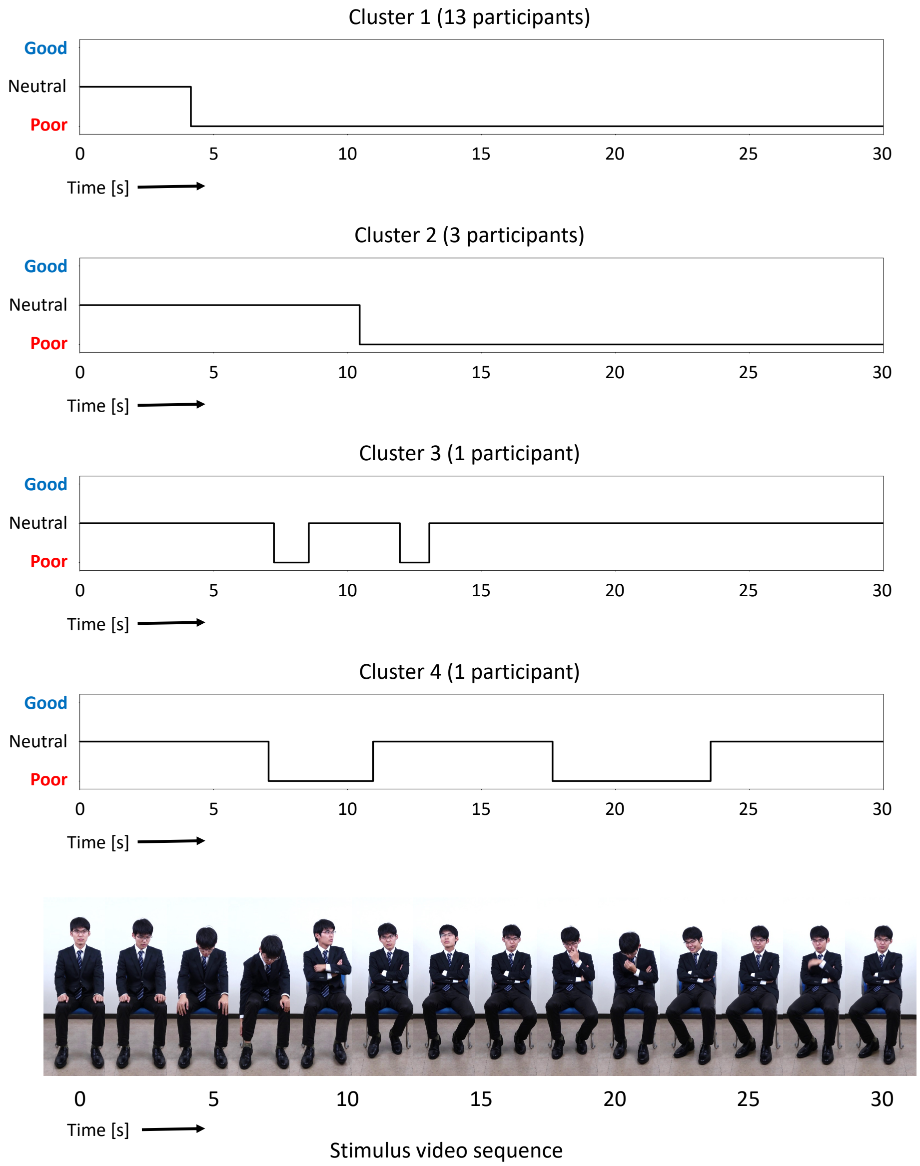

Fig. 7. Clustering results of the temporal changes of impression for the stimulus video sequence with the *lowest* total subjective score in the sequential evaluation.

6 Conclusions

We have developed a system that records the temporal changes of impression that interviewers form from the applicant's behavior during viewing of a stimulus video sequence, and investigated inter-rater reliability across participants who simulated interviewers. The experimental results confirmed statistically significant moderate inter-rater reliability across participants. In addition, participants who used our system indicated that they were able to report impressions without feeling a substantial burden.

In future work, we plan to analyze in detail the relationship between the temporal changes of impression and the applicant's behavior. We also plan to increase the diversity of participants, such as age groups, cultural background, and expertise, to improve the reliability of the findings.

References

1. Ambady, N., Rosenthal, R.: Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. Psychological Bulletin **111**(2), 256–274 (1992)
2. Barrick, M., Shaffer, J., DeGrassi, S.: What you see may not be what you get: Relationships among self-presentation tactics and ratings of interview and job performance. Journal of Applied Psychology **94**(6), 1394–1411 (2009)
3. Bindemann, M.: Scene and screen center bias early eye movements in scene viewing. Vision Research **50**(23), 2577–2587 (2010)
4. Cowie, R., Douglas-Cowie, E., Savvidou, S., McMahon, E., Sawey, M., Schröder, M.: 'FEELTRACE': An instrument for recording perceived emotion in real time. In: ITRW on Speech and Emotion. pp. 19–24 (2000)
5. DeGroot, T., Gooty, J.: Can nonverbal cues be used to make meaningful personality attributions in employment interviews? Journal of Business and Psychology **24**, 179–192 (2009)
6. Fourcade, A., Malandrone, F., Roellecke, L., Ciston, A., de Mooij, J., Villringer, A., Carletto, S., Gaebler, M.: AffectTracker: real-time continuous rating of affective experience in immersive virtual reality. Frontiers in Virtual Reality **6-2025**, 1–21 (2025)
7. Gifford, R., Ng, C., Wilkinson, M.: Nonverbal cues in the employment interview. Links between applicant qualities and interviewer judgments. Journal of Applied Psychology **70**(4), 729–736 (1985)
8. Koo, T.K., Li, M.Y.: A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine **15**(2), 155–163 (2016)
9. Levashina, J., Hartwell, C.J., Morgeson, F.P., Campion, M.A.: The structured employment interview: Narrative and quantitative review of the research literature. Personnel Psychology **67**(1), 241–293 (2014)
10. McGovern, T.V., Tinsley, H.E.: Interviewer evaluations of interviewee nonverbal behavior. Journal of Vocational Behavior **13**(2), 163–171 (1978)
11. Peeters, H., Lievens, F.: Verbal and nonverbal impression management tactics in behavior description and situational interviews. International Journal of Selection and Assessment **14**(3), 206–222 (2006)